MOBILE EXPERTS

EXPERT INSIGHT FOR IoT SUBSCRIBERS

Edge Computing and Private LTE:

We need convergence for URLLC to succeed

April 2019

Introduction

Some IoT market areas will be fairly simple: Enable the software for an LTE network to handle NB-IoT. Distribute some IoT devices. Turn it on. Ramp it up.

But the URLLC market will have a very different dynamic. For URLLC to work, the radio network will not be enough. Edge Computing and a different business model will also be needed for many businesses to take advantage of 5G IoT technology.

This Expert INSIGHT lays out the direction that 5G IoT is likely to take, as specific vertical markets adopt low-latency radio technology and address other limiting factors at the same time.

Latency

People are always talking about latency in the 5G radio network. But it's not as simple as saying that the radio network supports a 1 ms latency. Some applications involve the transfer of significant data, or significant processing… and that takes time! The IoT application generally will need to interface with a Cloud service, and today the Cloud data centers that support AWS, Microsoft Azure, and Google Cloud Platform (GCP) are typically located hundreds of miles from the end users.

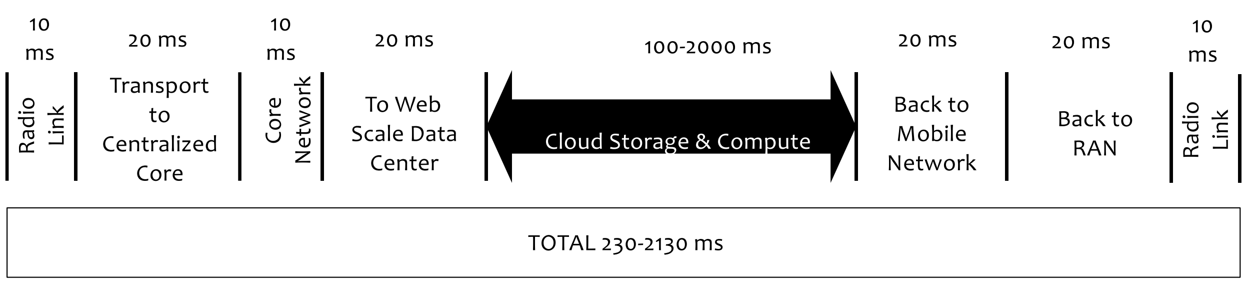

Figure 1: Latency is a lot more than the radio latency

Sources: Mobile Experts

Looking at the situation today, about 20-50 ms of latency can be saved in the radio layer in the upgrade from LTE to 5G. But another 200-2000 ms of latency must be saved in other areas in order to achieve the expected performance.

One big impact will come from better localization of the core network and the cloud service. The mobile operator needs to bring the core network (charging, policy control, authentication, etc) out to a local site, so that long-distance transport can be eliminated from the latency equation. Waiting 20+ ms for each long-distance hop can add up quickly.

Next, the biggest impact is likely to come in the cloud computing itself. The computing resources must be adequate to handle the task promptly. For example, in the Automotive market an HD map update may need to happen quickly. Let's assume that a crate fell off a truck, and of course the crate is not listed in the HD map as an obstacle for self-driving cars. The first car to the location will take a high-resolution photo and will need to upload significant data to the Cloud. The image must be analyzed and a decision must be made regarding how to update the HD maps. Then some information about a road hazard will be sent to any cars following on that road.

So, there is a balance between having local resources and having enough resources to perform the computing task quickly. Placing huge computing power at every tower site is unreasonable…so the industry will explore compromises where the operators work on a regional level to get the network and transport latency down into the 10 ms range…and then the Web scale guys invest heavily to get the computing function down to 10 ms or so.

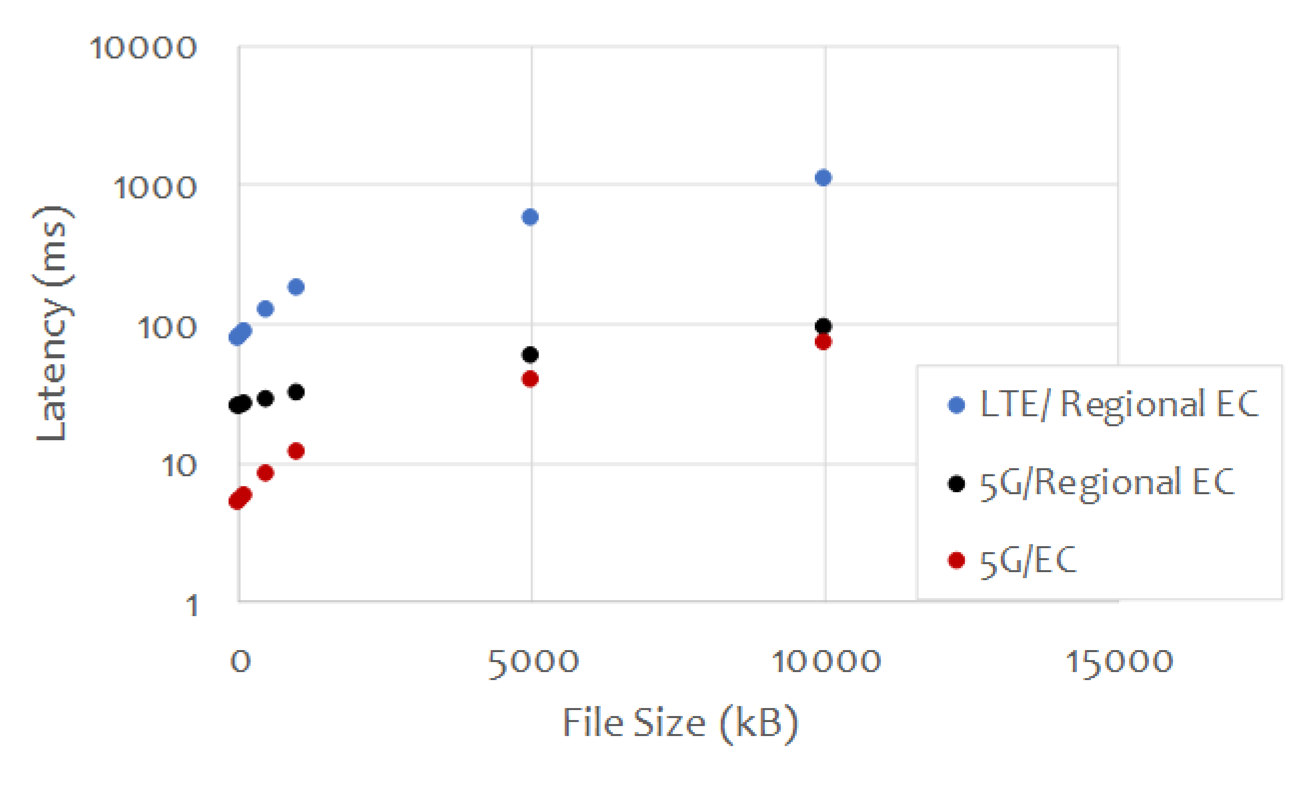

Figure 2: Message size is an important factor in the drive for low latency

Sources: Mobile Experts

Latency also depends on the size of the messages sent through the network. A big high-resolution image (say 1 MB) could take between 11 ms and 50 ms to transmit over the air, even for a 5G network with high nominal peak speed. Even the 'network slicing' function of 5G does not prevent the simple math of a 300 Mbps link trying to push a 1 MB file quickly.

In summary, the industry has been too simplistic to say that radio latency in the 1-2 ms range will create various IoT markets. In fact, the Edge Computing investment is also critical to achieving useful latency in real applications.

Business Models—Public Network

The enterprise applications for Edge Computing are starting to become more clear. We see them now, with facial recognition in airports, and license-plate readers, and other computing applications that have moved to the Cloud.

So far, the Cloud players have been happy to invest in regional colocation data centers, focusing their investment on the larger cities and specific applications that they want to target. But as the IoT and Edge Computing applications become a mainstream market, the best implementation will involve data centers in every city. Too big a challenge for a company like AWS to build their own, and the colocation specialists may not cover every location.

The mobile operators have a unique position here because mobile operators have real estate assets and operations everywhere. They've already established a footprint in every town that has more than 1000 people, and also in many places with far fewer people.

As a result, one business model is likely to emerge where the mobile operators become a host for Cloud Computing resources, and even charge rent to companies like AWS and Microsoft for space in the Central Office.

Business Models—Private Networks

Another major opportunity will come from Private LTE/5G networks. Large enterprises in several important markets have stated a strong desire to own their infrastructure, and control completely the data on the network. These companies don't want the involvement of the mobile operators:

- Electric, gas, and water utilities;
- Manufacturing operations;
- Oil and Gas operations;
- Transportation Networks; and
- Financial Institutions.

Because these companies want to control their data, in essence they require that the entire IoT network needs to live inside their firewall. That means that the IoT device, the radio network, core network, and Edge Computing resources will all be located inside the customer's building, and analysis of the data will be controlled by the company IT department as well.

In this case, the mobile operator can play in the game, if they agree to set up the IoT network within the facility and hand over the keys to the enterprise. The operator has some leverage in this situation, because the operators control most of the spectrum, and they have the expertise in most places.

But the operators will need to compete with Private LTE/5G networks that take advantage of other spectrum assets. The new CBRS band in the USA is one example. Germany's 3.7 to 3.8 GHz band is another example. As spectrum becomes more available to private enterprises, we expect to see more deployment of IoT networks that are completely independent of the mobile operators. (They'll probably still use NB-IoT, Cat-M, and 5G IoT technologies, but will not involve the operator in the business model).

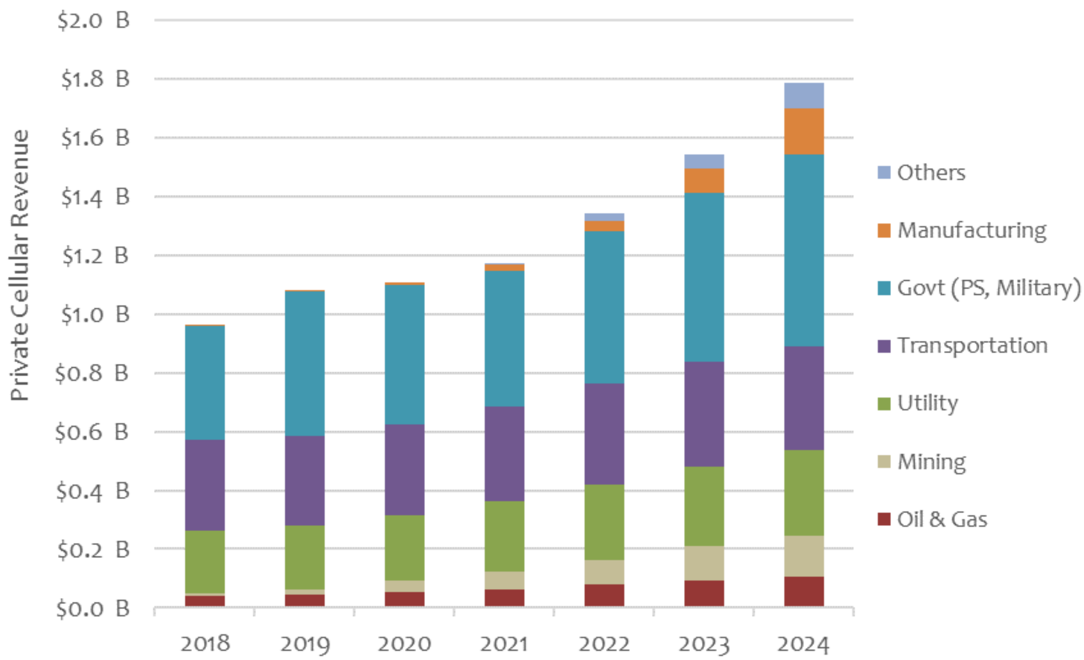

Figure 3: Growth in the Private LTE/5G market

Sources: Mobile Experts

So far, this market is pretty small. But Mobile Experts is predicting significant growth in various vertical markets.

1. Manufacturing will start from practically zero to become a significant market.
2. Government (Public Safety) is a well established market that will develop its own IoT applications such as fire monitoring devices, drones, and asset tracking.
3. Transportation has been strong for years (GSM-R and LTE-R are prime examples).

4. Utilities, Mining, and Energy markets will grow dramatically with the need for wide-area IoT devices. Each of these verticals will have company policies or regulatory requirements to maintain very tight security.

Conclusion

Creating a low-latency radio format is not enough. The business model must change, and the computing resources for many applications must be moved to a new location. In our view, all three of these major changes must come together in the market, before 5G URLLC IoT business can thrive.

Of course, each vertical market will move at a different pace. Each application will have different computing requirements, which drive distinct decisions with regard to edge computing placement. We believe that mobile operators will make 5G IoT services available over the next five years, but we will see larger growth in the 2025-2029 timeframe, when the Edge Computing and Private LTE investment models converge.